חפש מי ימצא אותך

מה זה ספיידר, מתי קטגוריות עדיפות, איך תומאס אדיסון קשור לזה, ואיפה לעזאזל מתחבאות כל הבחורות העירומות

עידו אמין

כבר אלף פעמים אמרו לכם את זה. באינטרנט יש הכל. תווים לגיטרה, חדשות מעודכנות, חברים חדשים, פילוסופיה ותמונות של בחורות עירומות. הכל. או.קיי. אז קמתם ועשיתם מעשה. התחברתם לאינטרנט, פתחתם את הדפדפן ו... עכשיו אתם שואלים: איפה לעזאזל מוצאים כאן הכל? לפני שאתם מחליטים שהאינטרנט זו התרמית הגדולה של המאה ה־20, קנוניה מטורפת של יצרני תוכנה ואזרחים טובי לב שלא נעים להם להודות שהמלך עירום, מוגש לכם בזאת האל"ף־בי"ת של חיפוש ברשת.

החיפוש באינטרנט לא קל, הדבר לא מזכיר אפילו חיפוש בארכיון או בספרייה, משום שכאן המידע שבו מחפשים מגובב ונובע ממקורות שונים בתכלית: עבודות אקדמיות, עלוני פרסום, קטלוגים של כל דבר שניתן לדמיין, מכתבים אישיים, אוספים של כרטיסי "חבורת הזבל" וזיכרונות מהתיכון. אם זה דומה למשהו, זה לחוויה של חיפוש בתוך קופסאות קרטון, רגע לפני שמסיימים לארוז את הדירה הישנה, לפני שהסבלים מגיעים.

זה לא כל כך גרוע כמו שזה נשמע. יש כמה וכמה מדריכים ומנועי חיפוש שיעזרו לכם לחפש. אל תרגישו בודדים, על פי סקר טרי של חברת נילסן, 71% מגולשי האינטרנט מחפשים בעזרת מדריך אתרים או מנוע חיפוש. פחות מ־10% גולשים לאן שהחברים המליצו, ועוד פחות על ידי גלישה מקישורים מאתרים אחרים, טלוויזיה ומדריכים מודפסים.

הרובוט עובד גם בלילה

גם הגולשים שעברו את הטירונות הווירטואלית עדיין נבוכים מכך שגם במנועי החיפוש המדהימים ביותר קשה מאוד למצוא בערימת השחת מחט, שאולי מסתתרת שם (ואולי לא). ראשית חוכמה – יש להבדיל בין ה"רובוטים" לבין המדריכים.

רובוט, (או "ספיידר", או "תולעת"), אינו באמת מכשיר הבנוי ממתכת, אלא תוכנה. הרובוט יוצא עם לילה, גולש באינטרנט, ומביא הביתה (לאתר החיפוש, למשל אלטוויסטה או לייקוס) ערימות של דפי מידע שבהם נתקל בדרכו. לעתים הוא גולש אל אתרים שעליהם נודע לו מפי בעליהם, ומשם הוא ממשיך לקישורים המופיעים בהם. קווי הקישורים שבין אתר לאתר באינטרנט הם מסלולי ההליכה שלו, ומדי פעם הוא חוזר ומבקר באותם אתרים כדי לראות מה השתנה.

מטרתם של אתרים מבוססי רובוטים: ליצור על הדיסק שלהם העתק מושלם ככל האפשר של עיקרי המידע באינטרנט. הם מביטים בכותרות העמודים, מנתחים את תוכנם, מחפשים בהם סימני דרך שמנהלי האתרים משאירים בהם, ונמנעים מלחטט בעמודים שסומנו על ידי בני אדם כ"מחוץ לתחום לרובוטים".

עד כאן התיאוריה. וכעת לדוגמה. נניח שמה שאתם ממש צריכים כרגע זה אינפורמציה חמה על תומאס אדיסון, ממציא הגרמופון. נבקש ממנוע החיפוש לאתר עבורנו עמודים שבהם מוזכרות המלים "אדיסון" ו"גרמופון". החיפוש יניב, קרוב לוודאי, יותר מדי תוצאות. לכן, רצוי להמשיך ולהוסיף מלים לשורת החיפוש, על מנת לצמצם את מגוון העמודים המופיעים. צריך לזכור: ככל שנוסיף מושגים רבים יותר, יתחיל המנוע להפעיל "שיקול דעת" משלו, ויביא לנו עמודים שבהם נזכרות רק חלק מהמלים.

באלטוויסטה, למשל, ניתן להוסיף סימן הפלוס ("+") לפני מלה, ואז מלה זו תהיה חייבת להופיע בעמודים המוגשים לנו. סימן מינוס ("-") יביא לסילוק העמודים המכילים אותה מן התוצאות.

עוד טיפים: צירוף מלים שהוא ביטוי, יש להקיף בדרך כלל במירכאות כדי שהמכונה תבין שיש לחפשן יחד. ברוב מנועי החיפוש ניתן לכתוב שאילתות בשפה מורכבת יותר, תוך שימוש במושגים לוגיים כגון And ו־Or. יש מנועים שיביאו לנו רק תמונות שתיאוריהן או שמותיהן יכללו את המילים המבוקשות, או יגבילו את חיפושיהם על פי מיקום העמוד, השרת שלו ועוד ועוד.

זה קשה, זה מסובך ולמי יש כוח. אבל יש גם אלטרנטיוות: מדריכים מסווגים.

שובה של האינטליגנציה האנושית

מדריך מסווג (למשל, יאהו!) אינו משתמש כלל באינטליגנציה מלאכותית, אלא באינטליגנציה אנושית. מטרתו העיקרית של מדריך מסווג היא ליצור מאגר מצומצם יותר ואיכותי יותר. לכן, אתרים חדשים הנרשמים במדריך מסווג אינם מופיעים מיידית במדריך, אלא לאחר שביקר בהם סוקר אנושי, עיין בתיאורם כפי שניתן על ידי בעליהם ואולי כתב תיאור קצר משלו. הסוקר מסווג את האתר באחד או יותר מהסיווגים ("קטגוריות") הנמצאים אצלו. במקרה של יאהו!, עץ הסיווגים מרשים מאוד.

עץ הסיווגים מאפשר שימוש במדריך על ידי דפדוף בסיווגים שבו. דרך נוספת לשימוש במדריך מסווג היא לחפש בעזרת מלות מפתח. אך, שוב, יש הבדל בין החיפוש במדריך מסווג לבין מנוע רובוטי: במדריך מסווג החיפוש נעשה בתיאורי האתרים המופיעים במדריך, ולא בתוכן העמודים עצמם. ייתכן שהאתר המוצלח ביותר למציאת חומר על אדיסון והגרמופון הוא "ההיסטוריה של הוויניל – תולדות הממציאים בעולם המוסיקה הפופולרית", שם שאינו מכיל לא "אדיסון" ולא "גרמופון".

הגולשים בעולם מעדיפים, כנראה, את שיטת המדריך המסווג. ההצלחה המסחרית של יאהו! משאירה את בעלי מנועי החיפוש האוטומטיים אכולי קנאה, ולפי הדיווחים האחרונים – גם עם קצת בעיות כלכליות.

מה עדיף? כאשר ידועים לנו השמות, או כשברורה הקטגוריה, כדאי להתחיל ולחפש במדריך מסווג. לעומת זאת, כאשר ניתן לחפש מלים מדויקות מאוד, עדיף כמובן לחפש במנוע רובוטי. ביאהו! כרתו ברית עם אלטוויסטה, וכאשר המלים אינן מופיעות ביאהו!, מועבר הגולש אוטומטית לאלטוויסטה לתרצאות החיפוש של אותן מלים. אם לא הצלחתם לדוג ולו אתר אחד במנוע חיפוש חופשי, נסו להזין את שמו או כתובתו למנוע מסווג כיאהו!, ותמצאו באיזו קטגוריה הוא סווג שם. זו בעצם תשובה מהירה לשאלה "מה עוד יש כמו זה", ותרגיל זה יכול גם להניב תוצאות יפות וגם ללמד רבות על מבנה מערכת הסיווגים.

כשמדובר באתרים ישראליים, התשובה מסובכת עוד יותר. לגבי יאהו!, התשובה פשוטה: אתרים ישראליים מסווגים בראש ובראשונה בסיווג "מקומות/מדינות/ישראל", שמתחתיה היירכיה של סיווגים שונים. אבל כשמדובר בסיווגים – עדיפים מדריכים בשפת האם, ואם ברצונכם לחפש אתרים ישראליים, תוכלו להשתמש בכמה מדריכים מסווגים בעברית. האתרים הישראליים נקלטים, כמובן, גם על ידי מנועים מבוססי רובוטים. הם אינם עולים בדרך כלל בתוצאות החיפוש, בעיקר משום שהחיפוש נעשה אחר מלים לועזיות שאינן מופיעות בעמודים אלה, אבל הם קיימים במסדי הנתונים של אלטוויסטה ודומיו. נסו פעם לחפש שם בעברית, "הפוכה" או לא, ותופתעו מהתוצאות.



איור: ט.ל.מ.

מנועי החיפוש הגדולים בעולם:

- אלטוויסטה http://altavista.digital.com
- אקסייט http://www.excite.com
- הוט־בוט http://www.hotbot.com
- אינפוסיק http://www.infoseek.com
- לייקוס http://www.lycos.com
- ווב־קרולר http://www.webcrawler.com
- אופן טקסט http://www.opentext.com
- יאהו! http://www.yahoo.com

מדריכים ישראליים:

- המדריך הכמעט שלם ל־WWW בישראל – נערך מאז 1994 על ידי סטודנט מהטכניון, מכיל מעל 1,400 אתרים
http://www.math.technion.ac.il/~nyh/israel
- ואללה! – מדריך מסווג גדול לאתרים עבריים (ולועזיים נבחרים), הכולל מנוע חיפוש פנימי וצ'ט בעברית
http://www.walla.co.il
- תפוז – מדריך מסווג הכולל כ־900 אתרים בעברית (ועוד כ־100 באנגלית), כולל גם מנוע חיפוש פנימי
http://www.tapuz.co.il
- דפי רשת – מדריך מסווג עם יותר מ־1,500 תיאורים של אתרים ישראליים, ומנוע חיפוש פנימי
http://dapey.reshet.co.il
- טופ טן – 10% האתרים הישראליים המוצלחים ביותר
http://www.top10.co.il
- יאהו!: אתרים ישראליים
http://www.yahoo.com/Regional/Countries/Israel
- סיבוב – כתבו מלה בעברית, היא תתורגם לאנגלית ותועבר מיד למנועי החיפוש הגדולים בעולם
http://www.sivuv.co.il/index.html